U. S. NAVAL SUBMARINE MEDICAL CENTER

Submarine Base, Groton, Conn.

REPORT NUMBER 594

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS BALANCING IN THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

by

James F. Willott, Cecil K. Myers
and
J. Donald Harris

Bureau of Medicine and Surgery, Navy Department
Research Work Unit MF12.524.004-9010D.06

Released by:

J. E. Stark, CAPT MC USN
COMMANDING OFFICER
Naval Submarine Medical Center

2 September 1969

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS BALANCING IN THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

by

James F. Willott, Cecil K. Myers
and
J. Donald Harris

SUBMARINE MEDICAL RESEARCH LABORATORY
NAVAL SUBMARINE MEDICAL CENTER REPORT NO. 594

Bureau of Medicine and Surgery, Navy Department
Research Work Unit MF12.524.004-9010D.06

Transmitted by
J. Donald Harris, Ph.D
Head, Auditory Research Branch

Reviewed and Approved by:

Charles F. Gell

Charles F. Gell, M.D., D.Sc.(Med)
Scientific Director
SubMedResLab

Reviewed and Approved by:

Joseph D. Bloom.

Joseph D. Bloom, CDR MC USN
Director
SubMedResLab

Approved and Released by:

J. E. Stark

J. E. Stark, CAPT MC USN
Commanding Officer
Naval Submarine Medical Center

SUMMARY PAGE

THE PROBLEM

To determine the sources of variance, and their extent, in the traditional psychoacoustic method of determining the real-ear response of an earphone when coupled to actual human heads.

FINDINGS

Variance due to (a) physical coupling of earphones and heads was estimated ab ± 1 dB, (b) collecting absolute thresholds was estimated at ± 2.25 — 3.5 dB, and (c) collecting alternate interaural loudness equality judgments was determined to be ± 1.5 — 2.5 dB. It is the totality of these variances which is inherent in the imprecision of the traditional procedure. The total variability involved in comparing a standard earphone with a new earphone was determined to be ± 4.16 — 7.54 dB for different individuals, considered too large for acceptable precision. One would have to recommend the massing of subjects for such comparisons, or preferably the development of new psychoacoustic procedures.

APPLICATION

For electrical engineers, sonar technicians, communications engineers, otologists, audiologists, and others interested in the specification of the real-ear response of an earphone.

ADMINISTRATIVE INFORMATION

This investigation was conducted under Bureau of Medicine and Surgery Research Work Unit MF12.524.004-9010D—Optimization of Auditory Performance in Submarines. The report has been designated as Submarine Medical Research Laboratory Report Number 594. It is report No. 6 on this work unit and was approved for publication on 2 September 1969.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL CENTER

ABSTRACT

The variances were examined associated with the psychoacoustic calibration of audiometric earphones by the usual method of alternate interaural loudness balancing with a standard earphone first on one side of the head and then on the other. Eight major sources of variance can be identified, the result of coupling two earphones first on one ear and then on the other, collecting two absolute thresholds first on one ear and then on the other using the standard earphone, and collecting two interaural loudness balances. On 13 subjects the differential sensitivity for alternate interaural loudness balancing was 1.5 — 2.5 dB, the higher frequencies giving somewhat larger values. Variances due to coupling and to absolute threshold testing were estimated at about 1 dB for the former and from 1.27 — 2.51 dB for the latter. It was considered that the sizes of these variances were quite sufficient to explain the test-retest consistency of mean transfer functions, which were of the order of 6 dB. The group mean transfer function could be specified for either of two new circumaural earphones with a precision of 0.92 — 4.84 dB at various audiometric frequencies ($\pm$ 1 Standard Error).

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS BALANCING IN THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

INTRODUCTION

In transferring audiometric threshold sound pressure levels from a standard earphone to an earphone of different sensitivity and physical configuration it is sometimes unsatisfactory to measure levels generated by the two earphones successively in a closed acoustic coupler such as a small metal cavity, or even in the actual cavity enclosed by an earphone placed on a human head. Where this is the case, as with an insert earphone or one of the large circumaural earphone/cushion units, a psychoacoustic loudness balance must be performed with a panel of listeners making loudness equality judgments at each audiometric frequency between a standard and a new earphone. For example, the ISO 1964 specification[1] for reference equivalent threshold sound pressure levels, given for five different earphones from five different countries, is based partly on psychoacoustic judgments of loudness equality performed by one or more laboratories in each country on the standard phone in that country compared with that of two other countries (see Weissler[2]).

Differences amounting to several decibels are seen between the mean transfer functions from different laboratories, relating the voltage on one phone to the voltage on another type of phone which yields equal loudness. The only recourse one has is to mass observers and observations until the standard error of the mean transfer function is acceptably small; but this approach is costly in time, and short-cuts are often adopted, to the degradation of the data.

The more exact statement of the sources of variances which combine in a total earphone transfer function has never been made, and little quantitative data are at hand on the extent of the variance ascribable to each source. This paper is an attempt at such a statement for two sources, (a) the variability involved in making a series of alternate interaural loudness balances, and (2) the total reliability of the transfer function for the individual as well as for the group.

METHOD

Subjects. Twelve subjects were graduate students in sensory psychophysiology, all with essentially normal hearing from 250 — 8000 Hz. One was an older experienced psychoacoustician with a mild high-tone hearing loss. All were very experienced in making both absolute and differential auditory judgments.

Apparatus. The two channels of an Allison Model 21B audiometer were used; each channel led to an impedance-matching network, 1-dB/step attenuator, and earphone. A Western Electric 705A earphone was the standard, the unknown a Permoflux PDR-600 driver encased in one of two circumaural earphone cushions, the Maice "Auraldome" and the TRACOR Corp. "Otocup."

Procedure. The experimenter seated the subject in front of the audiometer and took absolute thresholds at each frequency by the usual Method of Limits in 1-dB steps. Then the voltage was raised 40 dB, and the subject was asked to manipulate, at will, the two channel-interrupter switches (normally "OFF") on the audiometer, and the gain control of the unknown phone (without, of course, any visual cue) until he could report loudness equality between the two ears. Without moving the phones, ten such judgments were demanded. The phones could then be removed at will before a subsequent set of ten judgments. The usual counterbalancing of frequencies, ear order on the standard phone, and circumaural earphones was accomplished to allow for fatigue, order effects, etc.

TABLE I

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS DISCRIMINATION

Entry: Standard Deviation in DB of Ten Consecutive Loudness Equality Judgments at 40 DB Sensation Level on Standard Earphone. Comparison Earphone: Maico Co. "Auraldome"

Subj.	Stand. Phone On	Frequency in KHz								
		0.25	0.50	0.75	1	2	3	4	6	8
AR	L	3.26	1.89	1.96	1.96	3.26	2.04	2.97	3.29	3.26
	R	.83	2.00	1.73	2.10	2.87	2.11	3.61	2.32	2.28
JR	L	1.22	2.02	1.04	1.68	1.48	2.65	3.26	1.55	3.22
	R	.77	1.60	1.17	1.94	1.00	1.62	1.64	1.68	2.83
JD	L	1.99	.92	2.04	1.04	1.80	1.22	1.28	1.20	.98
	R	1.27	1.04	1.17	1.87	1.95	2.36	1.91	1.37	2.19
JW	L	1.34	2.62	1.27	3.68	2.47	2.30	2.19	1.90	3.35
	R	1.55	1.11	1.37	1.18	1.17	1.87	3.07	3.03	3.23
HM	L	3.69	2.26	1.56	2.00	2.59	2.29	3.54	1.85	2.40
	R	1.11	1.64	2.71	2.74	3.72	1.95	4.85	4.05	2.18
CMc	L	.66	.64	1.37	1.20	1.17	1.58	2.27	2.33	3.46
	R	1.19	.81	1.44	1.25	2.00	2.38	1.47	1.63	1.20
MD	L	2.34	1.47	2.11	2.47	2.19	1.40	3.21	2.24	2.19
	R	1.36	1.80	1.40	2.29	2.29	2.73	5.62	4.45	4.38
DW	L	1.43	2.24	3.01	.92	.80	3.19	2.32	2.90	3.96
	R	1.20	2.76	2.37	2.15	2.33	2.49	3.23	2.68	3.07
EC	L	2.00	2.06	2.10	3.01	2.33	1.91	2.48	2.45	2.49
	R	1.80	2.42	1.83	3.75	1.64	3.10	2.88	3.87	3.25
CM	L	2.01	2.19	2.91	1.76	2.41	1.67	1.36	1.94	2.15
	R	1.70	1.69	2.15	1.51	1.85	1.76	1.80	1.48	2.78
MH	L	1.19	1.55	2.54	2.01	2.32	2.11	2.76	1.95	2.43
	R	1.14	1.19	1.37	2.06	1.63	1.12	1.33	.90	2.53
RC	L	2.41	2.37	2.46	2.10	3.19	1.94	3.46	2.16	2.33
	R	2.98	2.42	2.97	2.12	1.69	2.68	1.83	2.68	2.28
RG	L	1.18	2.27	1.54	1.14	1.27	1.70	1.70	2.00	3.08
	R	1.68	2.18	.67	.98	1.43	2.61	1.96	2.68	.98
Mid-	L	1.99	2.06	2.04	1.96	2.32	1.94	2.48	2.00	2.49
Score:	R	1.27	1.69	1.44	2.06	1.69	2.36	1.96	2.68	2.53
	L+R	1.39	1.84	1.64	2.00	1.90	2.07	2.40	2.20	2.51

RESULTS AND DISCUSSION

1. Precision of Alternate Interaural Loudness Discrimination.

Precision in this case can be assessed by the standard Deviation (SD) of the ten consecutive equality judgments, considered as the differential sensitivity for alternate interaural loudness discrimination. Tables I-II give the SDs for both earphones. There is no systematic difference between R and L ears, and the last rows give the mid-value of each set of SDs. These values increase progressively from about 1.5 at 250 to about 2.5 dB at 8000 Hz.

No data of just this type have come to our attention, but there exist several sets of data from simultaneous interaural (i.e., dichotic) loudness judgments (for reviews see Harris[3] and Rowland and Tobias[4]). The latter have provided mean sensitivities of 1.15, 0.72, and

0.92 dB at .25, 2, and 6 kHz respectively, as compared with values of 0.88, 0.65, and 0.93 dB for the monotic condition, at overall loudness comparable to ours. If there is an effect of frequency, it is negligible.

Unfortunately, no estimates of variance were included, so that the precision of such values cannot be estimated and compared with ours. Furthermore, their subjects tracked the presence of intensity modulation, in a modification of the Bekesy Method of Limits, so that the mean sensitivity for each subject was an average of judgments "just noticeable difference" and "just not noticeable difference"; these are more traditionally called "jnd" and "jnnd," and their average the JND. It cannot be compared directly with the differential threshold (DL) from the Method of Constants, nor to the SD from the Method of Adjustments without appropriate transfer studies, which have never been done completely for loudness discrimination. Thus, the means in Tables I-II are a function both of the underlying sensitivity and of the variant

TABLE II

DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS DISCRIMINATION

Entry: Standard Deviation in DB of Ten Consecutive Loudness Equality Judgments at 40 DB Sensation Level on Standard Earphone. Comparison Earphone: TRACOR Corp. "Otocup"

Subj.	Stand. Phone On	0.25	0.50	0.75	1	2	3	4	6	8
		Frequency in KHz								
AR	L	1.62	1.92	2.28	1.91	2.19	1.18	2.32	.70	1.79
	R	1.60	1.25	1.78	1.78	1.20	2.68	4.96	4.17	2.96
JR	R	1.60	1.25	1.78	1.78	1.20	2.68	4.96	4.17	2.96
	R	1.43	1.37	1.44	1.17	1.36	.98	1.33	.81	2.42
JD	L	2.65	1.20	1.70	2.29	1.99	2.68	2.15	2.91	2.02
	R	1.37	1.58	1.14	1.79	1.62	1.41	2.76	2.42	3.67
JW	L	.90	1.00	2.26	2.00	1.90	.94	.92	2.21	2.51
	R	2.66	.81	1.20	.46	.98	1.42	1.83	1.25	2.60
HM	L	1.55	1.43	1.43	2.97	2.49	3.32	3.16	2.05	2.29
	R	2.06	1.73	2.34	1.14	2.15	1.99	3.46	2.77	2.83
CMc	L	1.04	1.22	2.19	.94	1.28	1.43	2.16	1.60	3.00
	R	1.68	.98	1.19	1.60	1.72	2.09	1.96	1.97	2.06
MD	L	1.33	1.37	1.04	2.49	1.72	2.53	2.00	2.43	2.61
	R	1.20	2.00	1.56	1.81	2.19	2.28	2.68	2.76	2.57
DW	L	1.08	1.86	2.42	1.84	2.80	3.61	2.11	2.24	6.26
	R	2.33	2.14	1.66	2.83	2.96	2.15	3.83	3.26	4.49
EC	L	1.11	2.15	1.60	2.66	1.43	4.32	2.16	2.24	1.92
	R	1.47	1.95	2.30	1.74	1.10	1.83	1.02	1.33	2.24
CM	L	2.29	1.47	1.78	1.87	1.30	1.11	2.02	1.42	1.55
	R	1.33	1.64	1.56	1.68	1.49	1.19	1.86	1.20	2.00
MH	L	2.18	2.01	2.49	3.10	1.64	2.64	2.16	3.37	2.73
	R	2.99	2.77	2.40	1.80	1.50	3.27	2.90	2.29	3.47
RC	L	3.74	3.01	2.99	3.03	4.71	5.21	4.46	5.04	6.14
	R	3.78	4.85	3.74	4.67	4.44	4.27	3.80	3.52	3.01
RG	L	2.00	3.23	1.49	3.88	3.10	3.72	4.45	3.74	5.99
	R	1.80	1.86	2.84	1.43	2.53	4.92	2.38	2.43	.83
Mid-	L	1.55	1.47	1.78	2.29	1.90	2.53	2.15	2.24	2.51
Score:	R	1.68	1.73	1.66	1.74	1.62	2.09	2.68	2.42	2.60
	L+R	1.61	1.68	1.74	1.82	1.72	2.27	2.16	2.35	2.55

of psychophysical method, but the contribution to the mean sensitivity and to its variance cannot be estimated at this time.

It might be supposed that the simultaneous interaural judgment be more sensitive and less variable, since an additional cue is present, namely, directionality of the phantom image in phenomenonological space. A relatively slight dichotic difference in intensity might move the sound image left or right and act as a vernier on a coarser scale. A direct comparison of this possibility was made by Jerger and Harford,[5] who found no difference in sensitivities by the Method of Adjustment for normal-hearing persons, though with persons with unilateral hypacusis the picture was entirely different. In such persons, differences between the two methods of as much as 10 dB were common. Evidently, the two types of interaural judgment are not at all the same, though they do yield the same SDs on normal subjects. Jerger and Harford did not furnish any estimate of individual or group variance.

2. The Distribution for our Population of the Transfer Functions.

A look at the transfer function distributions tells one at a glance how each new phone is reacted to by the group. Table III gives the SDs of the individual transfer functions, and the standard errors of the mean transfer functions. It is seen that massing observers to the number of 13 yields a true mean of a function at any frequency with a precision of 0.92 — 4.84 dB (± 1 Standard Error).

3. Test-Retest Reliability of the Earphone Transfer Function.

The complete description of the sources of variance in the function relating voltage on one phone to the voltage on another phone at equal loudness must take into account (a) coupling effects, (b) the reliability of the absolute thresholds, and (c) the reliability of the type of loudness judgment demanded. Thus, one has as a bare minimum those variances associated with:

(1) (2) coupling of the standard phone to the L ear, and again to the R ear when the phones are reversed to account for ear acuity differences,

(3) (4) same for unknown phone on the opposite ear,

(5) (6) the constant, variable, and accidental errors inherent in the taking of absolute thresholds with the standard phone on the two ears, and

(7) (8) the error associated with the two sets of loudness judgments.

In the previous section we have seen that the SDs for (7) (8) are of the order of 1.5 — 2.5 dB. Studies too numerous to mention have considered the separate and total variance associated with (1) — (6). Concerning only (1) — (4), Harris[6] showed that the SD of ten consecutive threshold crossings at all octaves 256 — 8192 Hz deteriorated by no

TABLE III

INDIVIDUAL DIFFERENCES IN THE EARPHONE TRANSFER FUNCTION, AND THE PRECISION OF THE MEAN TRANSFER FUNCTION

Entries: Standard Deviations of Distributions of Individual Transfer Functions, and Standard Errors of the Mean Functions

	Frequency in KHz								
	0.25	0.50	0.75	1	2	3	4	6	8
				"Otocup"					
SD	2.23	1.58	1.98	2.36	3.69	3.54	4.14	5.66	8.37
$S.E._{Mn}$	0.64	0.46	0.57	0.68	1.07	1.02	1.20	1.64	2.42
				"Auraldome"					
SD	5.42	4.43	3.43	3.61	3.75	5.40	4.79	6.82	6.28
$S.E._{Mn}$	1.57	1.28	.99	1.04	1.08	1.56	1.38	1.97	1.82

TABLE IV

DATA ON DISTRIBUTION OF TEST-RETEST DIFFERENCES IN EARPHONE TRANSFER FUNCTIONS BY THE TRADITIONAL LOUDNESS-BALANCING PROCEDURE

Entries: Mean Differences for Individual Transfer Functions, and the Standard Deviation and Standard Error of each Mean Difference

	Frequency in KHz								
	0.25	0.50	0.75	1	2	3	4	6	8
Mean	6.30	6.30	6.46	5.76	5.22	4.16	4.38	7.54	7.16
SD	5.30	4.82	4.24	4.50	4.98	3.74	2.82	5.64	6.12
$S.E._{Mn}$	1.47	1.34	1.18	1.47	1.38	1.04	0.78	1.56	1.70

more than about 1 dB when a comparison was made between the condition of removing-replacing, or not, the earphone after each threshold crossing. Hickling[7] found deterioration of SDs under these conditions of 0.5, 0, 1.35, and 1.32 dB at 1, 2, 6, and 8 kHz respectively. A conservative estimate of the contribution of each of the factors (1)—(4) to unreliability of the transfer function would be about 1 dB with experienced subjects. Of course, with some types of earphone the fit to the head might be more critical, or more difficult to standardize, and the variance might be larger.

The variance associated with (5) (6), including in most cases also (1)—(4), has often been assessed (for a review of ten such sets of data see Hickling[7]). The latter gave test-retest audiometry to 60 adults and computed intra-subject SDs of 2.27, 2.23, 3.51, and 3.40 dB at 1, 2, 6, and 8 kHz respectively. These figures are representatives of the deviations one may expect in establishing the thresholds of (5) (6).

The reader will see at once that it is not reasonable to add up the separate SDs associated with each of the eight sources, and represent that figure as the real variance of the total transfer function: these variances are all in terms of a $\pm$ sign, and would for the most part tend to cancel each other. Only an actual replication of the total process will give a realistic estimate of the total variance in the transfer function. What the estimates of factors (1)—(8) individually show is some basis for the variance which the total procedure actually yields.

In the present data an estimate of test-retest reliability was to be had. Test-retest audiometric threshold differences within either circumaural cushion used are no greater than those between cushions when corrected for differences in the sensitivity of the two drivers. Differences were therefore computed between the individual transfer functions from one circumaural earphone to the other. Data on these differences are in Table IV; they are given as the mean of individual test-retest differences. It is seen that the average difference will be as large as 6 dB at about half of the frequencies at random, with standard errors of 0.78—1.70 achieved. However, note that if a standard error at any frequency of no greater than 1.0 dB were desired, as should not be rare, the N would have to be raised to about 35 persons. This is hardly satisfactory, and ways should be sought to reduce the number or extent of the eight sources of variance inherent in this traditional procedure.

REFERENCES

1. Anon. ISO Recommendation R389, "Standard Reference Zero for the Calibration of Pure-Tone Audiometers." International Standards Organization, Nov. 1964. Available through the USA Standards Institute.
2. Pearl G. Weissler, "International Standard Reference Zero for Audiometers." J. Acoust. Soc. Amer., 1968, 44, 264-275.
3. J. D. Harris, "Loudness Discrimination." J. Sp. Hear. Dis., Suppl. 11, 1963.
4. R. C. Rowland and J. V. Tobias, "Interaural Intensity Difference Limen." J. Sp. Hear. Res., 1967, 10, 745-756.
5. J. F. Jerger and E. R. Harford, "Alternate and Simultaneous Binaural Balancing of Pure Tones." J. Sp. Hear. Res., 1960, 3, 15-30.
6. J. D. Harris, "Experiments on Fluctuations of

Auditory Acuity." **J. General Psychol., 1954, 50,** 87-109.

7. S. Hickling, "Studies on the Reliability of Auditory Threshold Values." **J. Audit. Res., 1966, 6,** 39-46.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) NAVAL SUBMARINE MEDICAL CENTER, Submarine Medical Research Laboratory	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE
DIFFERENTIAL SENSITIVITY FOR ALTERNATE INTERAURAL LOUDNESS BALANCING IN THE PSYCHOACOUSTIC CALIBRATION OF EARPHONES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Interim Report

5. AUTHOR(S) (First name, middle initial, last name)
James F. WILLOTT, Cecil K. MYERS and J. Donald HARRIS

6. REPORT DATE 2 September 1969	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS 7
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) SMRL Report Number 594	
b. PROJECT NO. MF12.524.004-9010D.06		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Submarine Medical Center Box 600 Naval Submarine Base Groton, Connecticut 06340

13. ABSTRACT

The variances were examined associated with the psychoacoustic calibration of audiometric earphones by the usual method of alternate interaural loudness balancing with a standard earphone first on one side of the head and then on the other. Eight major sources of variance can be identified, the result of coupling two earphones first on one ear and then on the other, collecting two absolute thresholds first on one ear and then on the other using the standard earphone, and collecting two interaural loudness balances. On 13 subjects the differential sensitivity for alternate interaural loudness balancing was 1.5 - 2.5 dB, the higher frequencies giving somewhat larger values. Variances due to coupling and to absolute threshold testing were estimated at about 1 dB for the former and from 1.27 - 2.51 dB for the latter. It was considered that the sizes of these variances were quite sufficient to explain the test-retest consistency of mean transfer functions, which were of the order of 6 dB. The group mean transfer function could be specified for either of two new circumaural earphones with a precision of 0.92 - 4.84 dB at various audiometric frequencies (± 1 Standard Error).

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Earphones, calibration of						
Loudness balancing in earphones						
Psychoacoustic calibration of earphones						